

# Attention, Perception, & Psychophysics

## Detecting single-target changes in multiple object tracking: The case of peripheral vision

Journal:	<i>Attention, Perception, &amp; Psychophysics</i>
Manuscript ID	PP-ORIG-15-140.R2
Manuscript Type:	Original Manuscript
Date Submitted by the Author:	n/a
Complete List of Authors:	Vater, Christian; University of Bern, Kredel, Ralf; University of Bern, Hossner, Ernst-Joachim; University of Bern,
Keywords:	Attention: object-based, eye movements and visual attention, visual search

SCHOLARONE™  
Manuscripts

Only

Running Head: DETECTING TARGET CHANGES WITH PERIPHERAL VISION

Detecting single-target changes in multiple object tracking:

The case of peripheral vision

Christian Vater<sup>1</sup>, Ralf Kredel<sup>1</sup>, & Ernst-Joachim Hossner<sup>1</sup>

<sup>1</sup>University of Bern, Switzerland

Corresponding author:

Christian Vater

University of Bern

Bremgartenstrasse 145

CH-3012 Bern

Switzerland

Tel: +41 31 631 52 07

Mail: Christian.Vater@ispw.unibe.ch

Names and mail addresses of co-authors:

Ralf Kredel, Ralf.Kredel@ispw.unibe.ch

Ernst-Joachim Hossner, Ernst.Hossner@ispw.unibe.ch

**Abstract**

In the current study it is investigated whether peripheral vision can be used to monitor multiple moving objects and to detect single-target changes. For this purpose, in Experiment 1, a modified MOT setup with a large projection and a constant-position centroid phase had to be checked first. Classical findings regarding the use of a virtual centroid to track multiple objects and the dependency of tracking accuracy on target speed could be successfully replicated. Thereafter, the main experimental variations regarding the manipulation of to-be-detected target changes could be introduced in Experiment 2. In addition to a button press used for the detection task, gaze behavior was assessed using an integrated eye-tracking system. The analysis of saccadic reaction times in relation to the motor response shows that peripheral vision is naturally used to detect motion and form changes in MOT because the saccade to the target occurred after target-change offset. Furthermore, for changes of comparable task difficulties, motion changes are detected better by peripheral vision than form changes. Findings indicate that capabilities of the visual system (e.g., visual acuity) affect change detection rates and that covert-attention processes may be affected by vision-related aspects like spatial uncertainty. Moreover, it is argued that a centroid-MOT strategy might reduce the amount of saccade-related costs and that eye-tracking seems to be generally valuable to test predictions derived from theories on MOT. Finally, implications for testing covert attention in applied settings are proposed.

**Keywords**

covert attention; perception; motor control; saccadic latency; eye-tracking, sports

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

47     **Introduction**

48             Peripheral vision allows us to detect natural objects at large eccentricities at up to  
49     70.5° eccentricity level (Thorpe, Gegenfurtner, Fabre-Thorpe, & Bülthoff, 2001) and to cate-  
50     gorize natural scenes at 70° eccentricity with high accuracy (Boucart, Moroni, Thibaut, Szaf-  
51     arczyk, & Greene, 2013) even under crowded conditions (Li, VanRullen, Koch, & Perona,  
52     2002). Neuroanatomically, the capability to process visual information in the far periphery  
53     seems to be controlled by distinct, specialized cortical networks (for a recent review see Yu,  
54     Chaplin, & Rosa, 2015). Furthermore, in comparison to foveal vision, peripheral vision fea-  
55     tures higher temporal resolution (Hartmann, Lachenmayr, & Brettel, 1979) and superior mo-  
56     tion detection (Finlay, 1982; McKee & Nakayama, 1984). Thus, while foveal vision with its  
57     high spatial resolution is important for extracting detailed information, a crucial role for pe-  
58     ripheral vision seems to refer to the processing of changing aspects in the visual environment  
59     (for a review, see Strasburger, Rentschler, & Jüttner, 2011).

60             Hence, a monitoring task on the detection of motion and form changes in the periphery  
61     can be regarded as a paradigmatic example of making use of peripheral vision. In an applied  
62     sense, those tasks can be found in team sports, where it is often necessary to perceive a num-  
63     ber of moving objects across a large functional field of view (Davids, 1984). This applies, for  
64     example, in offside decision-making in soccer, where the assistant referee has to monitor the  
65     player in ball possession as well as the players at the offside line. In this situation, experts  
66     show a gaze strategy in which they anchor their gaze on the offside line while perceiving the  
67     player in ball possession peripherally (Catteeuw, Helsen, Gilis, Van Roie, & Wagemans,  
68     2009). Furthermore, Williams and Davids (1998) were able to empirically show that, in a  
69     three-versus-three anticipation task in soccer, experts extract a great amount of information  
70     regarding players' positions and movements with peripheral vision. The importance of track-  
71     ing multiple objects in team sports was also underlined by Faubert and Sidebottom (2012),

who described the necessity for soccer goalkeepers to simultaneously track the ball, teammates, and opponents while keeping a large and dynamic visual field under surveillance.

In experimental psychology, the demands imposed from ecologically valid situations are generally studied with *multiple-object tracking* (MOT), a task, which was introduced by Pylyshyn and Storm (1988) to study visual attention. The task requires tracking a certain number of targets amidst identically looking distractors for a pre-defined amount of time. Before the objects begin to move on the computer screen, the targets are briefly highlighted, then all objects start to move in a (quasi-)random fashion, and finally, after targets have stopped, participants are supposed to recall the targets (see Figure 1 in which stimulus material of the current study is illustrated).

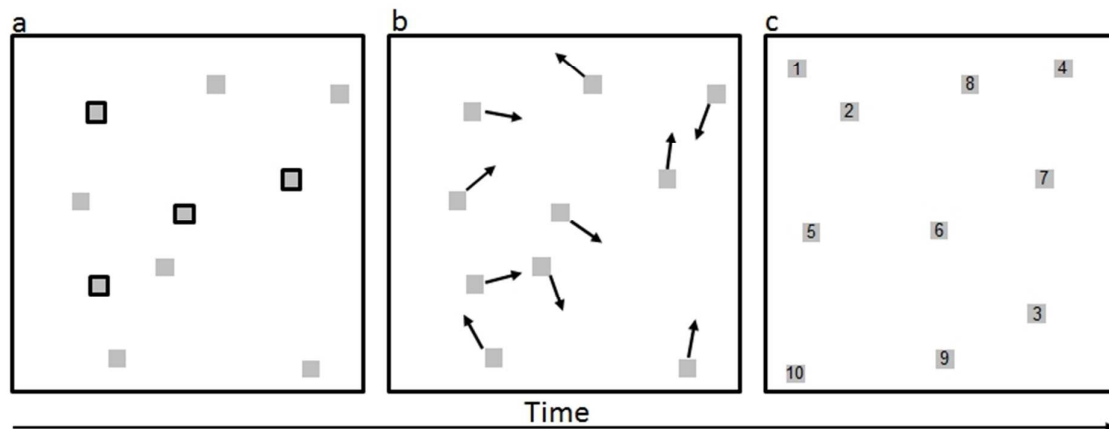


Figure 1. Experimental paradigm used in Experiment 1 and 2. After designating the four targets with red frames (a), the frames disappear and all stimuli move in quasi-random fashion for 6 s (b). At the end of the motion phase, participants have to recall the targets by naming the respective numbers that are projected onto the objects (c).

Theoretically, several approaches have been brought forward to explain tracking performance in MOT, namely, (a) the FINST (Fingers of INSTantiation) model by Pylyshyn and Storm (1988), (b) the grouping model by Yantis (1992), and (c) the multifocal attention model by Cavanagh and Alvarez (2005). After the brief descriptions of these approaches, we will

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

91 come back to the question how specifics of peripheral vision can be related to these concepts  
92 in particular and to MOT performance in general.

93 (a) In their MOT studies, Pylyshyn and Storm (1988) were mainly interested on how  
94 visual attention can be shifted independent of eye movements and on whether locations are  
95 scanned serially or in parallel. According to their feature-binding-based FINST model, partic-  
96 ipants track multiple objects in parallel while each object possesses an individual internal ref-  
97 erence at a pre-attentive stage. The authors describe a serial tracking algorithm in which target  
98 locations (instead of motion vectors) are stored. In a related experiment, motor response la-  
99 tencies to flashes on targets were measured as a function of the number of targets. Latencies  
100 significantly increased from one target to five targets which could be regarded as an indicator  
101 that participants serially track the targets (Pylyshyn & Storm, 1988). However, in a second  
102 experiment, to test the serial-scanning algorithm, participants had to indicate whether a flash  
103 occurred on a target, distractor, or another location. As the observed performance turned out  
104 to be much higher than the algorithmically predicted one, Pylyshyn and Storm finally con-  
105 cluded that targets are rather tracked in parallel than serially.

106 (b) Yantis (1992) alternatively proposed that participants keep track of the targets by  
107 making use of a virtual polygon formed out of the targets. Two stages are predicted by this  
108 model. The first one, the group formation stage, is influenced by Gestalt laws of grouping and  
109 can be characterized as pre-attentive, automatic, and stimulus-driven and, thereby, compara-  
110 ble to the model introduced by Pylyshyn and Storm (1988). However, in contrast to the  
111 FINST model, the second stage that concerns group maintenance is characterized as goal-  
112 directed, effortful, and attention-demanding. Furthermore, a continuous updating of target  
113 representations is assumed and discussed in the context of mental rotations. Yantis (1992)  
114 experimentally tested these predictions by applying different Gestalt laws to the MOT task.  
115 For example, when five targets either rotated in the same direction (rigid condition) or three

1  
2  
3 116 targets in one and two targets in the other direction (nonrigid condition), response accuracies  
4  
5 117 were higher in the rigid condition, presumably because targets followed the law of common  
6  
7 118 fate. Furthermore, Zhong Ma, Wilson, Liu, and Flombaum (2014) showed that, rather than  
8  
9 119 motion extrapolation, the recently observed object position seems to be used for tracking.  
10  
11 120 This finding can be interpreted as support for the effortful updating process in the Yantis  
12  
13 121 model because, on the basis of the continues updating, no motion extrapolation would be ex-  
14  
15 122 pected (for an overview on mixed results on motion extrapolation, however, see Zhong et al.,  
16  
17 123 2014).

18  
19  
20  
21 124 (c) Finally, Cavanagh and Alvarez (2005) proposed a multifocal attentional mecha-  
22  
23 125 nism for MOT that is based on the assumption that attention is split to the targets while each  
24  
25 126 selection channel comprises a position tracker and a stream of object features. This model is  
26  
27 127 underpinned, among others, by the results presented by Meyerhoff, Papenmeier, Jahn, and  
28  
29 128 Huff (2013) who were able to show that unexpected changes in targets do affect tracking per-  
30  
31 129 formance while changes in distractors have no effect even if these changes were in the line of  
32  
33 130 sight. To examine whether motion information are actually used for tracking, Huff and col-  
34  
35 131 leagues introduced motion-texture information on the targets moving either in the same or in a  
36  
37 132 different direction as a target, thereby, not confounding spatiotemporal information of the  
38  
39 133 objects with motion information (Huff & Papenmeier, 2013; Meyerhoff, Papenmeier, & Huff,  
40  
41 134 2013; St.Clair, Huff, & Seiffert, 2010). Among others, these studies show that tracking per-  
42  
43 135 formance declines when the difference between object direction and texture direction increas-  
44  
45 136 es (St.Clair et al., 2010) and that the availability and reliability of spatiotemporal and feature  
46  
47 137 information is considered by participants (Meyerhoff, Jahn & Huff, 2014). As motion and  
48  
49 138 feature information seem to be relevant for tracking performance, these results are perfectly in  
50  
51 139 line with the multifocal attention theory of MOT.  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

140 In the three models sketched so far, the theoretical focus was laid on questions on vis-  
141 ual attention, mostly without applying eye-tracking technology to the respective studies. In  
142 this vein, Fehd and Seiffert (2008) demonstrated that multiple target objects are tracked  
143 amongst distractors with a center-looking strategy relying on a virtual centroid, which is de-  
144 fined as the center of mass of the polygon formed by the targets. Furthermore, they showed  
145 that this strategy is independent of speed and object size and that tracking performance is  
146 highest when a center-target switching strategy is used in which gaze is shifted back and forth  
147 between the centroid and the targets. Most interestingly in regards to the observed gaze be-  
148 havior, slower stimuli did not lead to more saccades so that a saccade-avoiding strategy was  
149 not evident. On the one hand, this could have been expected since increased effects of crowd-  
150 ing are usually observed at higher stimulus speed (Franconeri, Lin, Pylyshyn, Fisher, & Enns,  
151 2008), which would lead to the necessity to saccade more frequently to targets because of  
152 potential collisions (Elfanagely Haladjian, Aks, Kourtev, & Pylyshyn, 2011; Landry, Sheri-  
153 dan, & Yufik, 2001). On the other hand, while saccading helps updating the position of single  
154 targets (Landry et al., 2001), the costs of eye movements increase with higher stimulus speed  
155 because, when executing saccades, objects move far if stimulus speed is high (Huff, Pa-  
156 penmeier, Jahn, & Hesse, 2010). Additionally, saccadic suppression of information pro-  
157 cessing (about 75 ms before saccade onset to 50 ms after saccade offset) makes the updating  
158 process even more difficult (Diamond, Ross, & Morrone, 2000). In this context, the role of  
159 “rescue saccades” in MOT has been examined by Zelinsky and Todor (2010) who found that  
160 anticipatory saccades are driven by the potential for losing track of a target when it is close to  
161 a barrier occluder or a distractor. It could, therefore, be that participants preferably use their  
162 peripheral vision to monitor target positions while “anchoring” their gaze on the virtual cen-  
163 troid (Fehd & Seiffert, 2010; Zelinsky & Neider, 2008) and initiate saccades just before the  
164 distance of a target to other objects becomes too small (Zelinsky & Todor, 2010).



When relating the attention-related explanations of MOT behavior (see above: a-c) to the just sketched gaze studies, it seems likely that covert attention is used for tracking as gaze is apparently on a virtual centroid in large part. Thus, to examine the actual location of attention, target changes has been included in experimental studies on MOT. In this branch of research, target movements (motion direction) or target properties (color, form) were manipulated to investigate, on the one hand, whether target-motion information are used for tracking and, on the other hand, how attention is distributed to targets and distractors. Regarding the use of motion information, Meyerhoff et al. (2013) studied the effect of random motion-direction changes (deviation of up to 60 ° to the left or right) of either targets, distractors, or both. The result, that an *unexpected* change of a target but not of a distractor vector impaired performance, shows that, in line with previous findings (e.g., Fencsik, Klieger, & Horowitz, 2007), target-motion information are crucial for MOT. However, if changes are *expected* and participants have to respond to them, it is possible to figure out where attention is located during MOT. Therefore, Bahrami (2003) introduced color (red, green, blue) and form (“T”, “L”, “+”) changes in targets and distractors. He found that color and form changes of the target were more often detected than changes of distractors, again showing that attention is attracted by targets. Moreover, color changes were identified more frequently (approx. 80 %) than form changes (approx. 55 %). In a closely related study, Sears and Pylyshyn (2000) induced either a target or a distractor form change in 50 % of the trials (shape of a “seven-segment box figure eight” becoming either an “E” or an “H”). They showed that target changes were detected much faster than distractor changes, while a higher number of distractors increased response times when target or distractor changes were evident, underlining the finding that attention is rather attracted by targets than by distractors. This result is also supported by electrophysiological measures, for example, by Drew, McCollough, Horowitz and Vogel (2009) who measured event-related potentials (ERPs) while participants had to track two targets amidst

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

four moving distractors and four stationary distractors. Task-irrelevant white square probes were presented randomly at the target, distractor, empty space or a stationary target. As expected, target probes showed the greatest ERP-response while weaker responses were observed for distractors and weakest for background and stationary targets. Hence, behavioral and electrophysiological results confirmed that spatial attention is more on targets than on other locations.

Summing up, regarding theoretical approaches to explain MOT performance, recent research rather supports the multifocal attention theory (c) as target motion and feature information are apparently used in the tracking process. Furthermore, the centroid strategy revealed by eye-tracking studies suggests that MOT performance is mainly based on covert attention which, in turn, is particularly allocated to the targets.

As it becomes obvious from this summary, despite the fact that eye-tracking technology has been applied, results revealed in MOT studies were mostly discussed from an attention perspective and rarely from a vision perspective. This comes as a surprise because many findings could also be explained by the basal properties of the peripheral visual system, in particular, by the decreasing visual acuity but increasing contrast sensitivity of the peripheral retina. For example, when Bahrami (2003) reports that target color changes are detected more often than form changes this could also be traced back to luminance and contrast sensitivity differences of the retina such that foveal vision might be advantageous in detecting form changes due to its higher spatial resolution (Gralla, 2007). Beside this, the higher temporal resolution (Hartmann et al., 1979) and the superior motion-detection capability of the peripheral visual system (Finlay, 1982; McKee & Nakayama, 1984) would also predict high detection rates for motion changes in MOT. In this regard, it is really surprising that, up to date, no attempt has been made to implement a motion-change detection task to examine the role of peripheral vision in MOT. In a respective study, it would seem crucial to control gaze behavior as in

previous experiments (e.g., Sears & Pylyshyn, 2000) participants were only instructed to track target objects while keeping their eyes on the fixation cross without including eye-tracking technology. Consequently, up to now, it is completely unknown which natural gaze strategy is executed in MOT to detect target changes in a setting in which no explicit instructions on vision are given. Hence, exactly this empirical question will be addressed in the following experiments.

From an applied perspective, it should be added that these experiments promise the derivation of helpful advices for the training of gaze strategies, for example, in team-sport situations as they were sketched in the beginning of this paper. In this regard, form-changes in MOT can be associated with posture-related changes (e.g., a basketball player indicating with his hands where he wants to receive the ball) and motion-changes in MOT with movement-related changes (e.g., detecting a penetration of an opposing player to the basket). As the reliability of such recommendations grows with the external validity of the experimental setting, it seems worthwhile to study those changes by displaying the MOT task on a large screen to overcome the limitation of a restricted field of vision when examining the relevance of peripheral vision. Hence, in a series of two studies, Experiment 1 aims to answer the question whether classical MOT findings can be replicated under those novel experimental conditions. Based on the results of Experiment 1, the main research question will be approached in Experiment 2 in which the focus will be on the case of peripheral vision in monitoring multiple moving objects while detecting changes in the periphery.

235

## 236 **Experiment 1**

Experiment 1 was designed as a replication study, in particular regarding the MOT results reported by Fehd and Seiffert (2008, 2010) on the effects of object speed on response accura-

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

cy, on the one hand, and on gaze behavior, on the other hand. To increase the external validity of the study, a large screen was used to display the MOT task. In addition, a novel manipulation detail was introduced in anticipation of the main research question on form versus motion changes in Experiment 2. In this experiment, it seems to be crucial to manipulate the motion of targets and distractors in such a way that form- or motion-related changes occur independently of (uncontrolled) movements of the centroid as well as of (uncontrolled) eccentricities of the to-be-detected event. For this reason, a stationary phase of the centroid was applied in which the respective event occurred at a defined eccentricity (under the assumption that the centroid was actually focused). In spite of the larger display and the introduction of a stationary-centroid phase, a successful replication of the findings reported by Fehd and Seiffert (2008, 2010) was expected. Hence, it was predicted, that response accuracy decreases as a function of object speed and that participants prefer a centroid-tracking strategy independent of speed conditions.

**Methods**

*Participants.* 14 sport science students (7 females and 7 males; age:  $21.7 \pm 1.3$  years) participated in the experiment and received course credits in return. They had self-reported normal or corrected-to-normal vision and were unaware of the research question. The experiment was undertaken in accordance with the Declaration of Helsinki.

*Stimuli.* Motion paths of the objects (10 white squares, 35 mm x 35 mm corresponding to  $1^\circ \times 1^\circ$  of visual angle) were calculated with Matlab and then imported to Autodesk 3ds Max (2014) to render single video trials. Stimuli were presented within a quadratic frame (white line of 25 mm width, 1.40 m x 1.40 m corresponding to  $40^\circ \times 40^\circ$  of visual angle) on a black background. All squares appeared in (quasi-)randomized starting positions precluding overlaps with each other. Each trial began with a trial number, followed by a stationary

263 presentation of ten squares, including four targets that were highlighted by red frames (line of  
264 15 mm width; frame and stimulus together covering an area of  $1.7^\circ \times 1.7^\circ$  of visual angle).  
265 After 2 s, target cues disappeared, and all stimuli accelerated over one second to the final  
266 speed of either  $6^\circ/\text{s}$  or  $9^\circ/\text{s}$  or  $12^\circ/\text{s}$ , which was retained for 4 s. After a subsequent decelera-  
267 tion phase of one further second, all squares stopped, so the overall motion time was 6 s per  
268 trial.

269 During the motion time, a repulsion mechanism was used to redirect a square whenever  
270 the distance to the bordering frame or to the next square fell below a certain threshold  
271 (35 mm corresponding to  $1^\circ$  of visual angle). To ensure independence of the findings of cen-  
272 troid-related motions (as well as of different eccentricities of target changes in Experiment 2),  
273 targets' paths were manipulated such that the centroid was forced to stay at a constant posi-  
274 tion for 0.5 s. For this purpose, the path of one distractor was calculated in such a way that it  
275 collided with one target at one of the three possible points in time (3.0 s, 3.5 s, 4.0 s after mo-  
276 tion onset) and that all targets moved after the collision in a concerted way in angles of  $90^\circ$   
277 difference to each other (e.g.,  $10^\circ$ ,  $100^\circ$ ,  $190^\circ$ , and  $280^\circ$  in relation to the frame's baseline).  
278 Due to the same speed of the targets, this specification resulted in a stationary position of the  
279 centroid of the polygon formed by the four targets. Over the 0.5 s following the critical colli-  
280 sion, the five remaining distractors were allowed to move randomly; however, it was ensured  
281 that no collision of a target with another object or the bordering frame occurred within this  
282 time window of 0.5 s. Before and after this interval, all objects were allowed to collide with  
283 one another or with the bordering frame without any restrictions.

284 On the basis of these constraints, 14 mother trials were created, which differed on the  
285 position of the crucial collision, the resulting position of the stationary centroid, and the con-  
286 certed motion direction of the target during the stationary-centroid phase. From the mother  
287 trials, a total of 126 trials ( $= 14 \times 3 \times 3$ ) was derived by starting the stationary-centroid phase

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

at three different onsets (either 3.0 s, 3.5 s, or 4.0 s after motion onset) and by accelerating targets and distractors to three final speeds (either 6 °/s, 9 °/s, or 12 °/s). To balance trials of different speeds within blocks, each block contained five trials of each speed, which were presented in a randomized order (rendering by MAGIX Video Pro X3). Since a single session of 60 min allowed for the presentation of a maximum of 9 blocks and participants' capability to keep focused on the quite demanding task seemed to be limited, exactly this number of 15-trial blocks was employed. Consequently, in the resulting total of 135 trials, 9 of the 126 trials derived from the mother trials were presented twice (i.e., three per stationary-centroid phase onset and speed).

*Apparatus.* A binocular eye-tracking system (EyeSeeCam, 220 Hz) was used to assess the vertical and horizontal rotations of both eyes via infrared reflections from the pupil and the cornea (measurement accuracy: 0.5 ° of the visual angle with a resolution of 0.01 ° RMS within 25 ° of the participant's field of view). The EyeSeeCam (ESC) is connected to a MacBook Pro via a 20 m fiber-optic Fire Wire link (GOF-Repeater 800, Unibrain), which is stored in a bum bag. Since the eye-tracker is synchronized with a 10-camera VICON-T20 system that tracks retro-reflective markers attached to the ESC, a three-dimensional gaze vector in a laboratory frame of reference can be calculated (Kredel et al., 2011). This gaze vector is updated every 5 ms and allows for relating the participant's current gaze to either the displayed targets or distractors or the centroid derived from the current positions of the targets.

The ESC was calibrated at the beginning of each test session. For this purpose, participants had to consecutively fixate five dots of a regular grid with a distance of 8.5 ° of visual angle between the dots (Kredel et al., 2011). The ESC was recalibrated before each test block whenever the point of gaze deviated more than 0.5 ° of visual angle from one of the dots in the calibration grid. Stimuli were back-projected (projector: InFocus IN5110) onto a large screen (height: 1.87 m; width: 3.01 m), while the quadratic frame for the MOT task covered

an area of 1.40 m x 1.40 m in the middle of the screen. Microsoft Windows Media Player was used to playback the video trials. The gathered data were analyzed using Mathworks MATLAB 2013a. Further statistical analyses were conducted with IBM SPSS Statistics 22.

*Procedure.* Participants were tested individually in a single one-hour session in the institute's sensorimotor laboratory. After reading the general information about the study and agreeing to participate, the eye-tracking system was fitted. Subsequently, participants were positioned at 2 m distance from the screen for reading the displayed instructions about the to-be-solved task. The task was to recall the targets that were cued at the beginning by naming the respective numbers that were projected onto the targets at the end of each trial. Afterwards, the ESC calibration routine was conducted. After each trial, participants had to recall the targets by naming the respective numbers that were displayed on the ten squares. Participants' verbal decisions were recorded in writing by an experimenter. No augmented feedback on the responses' correctness was given after the trials.

*Gaze analysis.* For the assessment of the location of gaze, a region-of-interest method was applied. For this purpose, eleven regions of interest were defined as circles with a diameter of  $5^\circ$  over each square as well as over the centroid, which was calculated as the center of mass of the polygon formed by the four targets (Fehd & Seiffert, 2010). Whenever the current gaze vector hit a region of interest, the respective frame was counted for the related object. This means that, in cases of overlapping regions, an allocation to two or even more objects was possible. The number of counted frames per region was then divided by the total number of frames in the trial (i.e., by 1200 frames for the 6 s of motion time) resulting in percentages of gaze-allocation time for each region of interest, a variable that has been termed "gaze overlap time" by Fehd and Seiffert (2010).



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

*Measures.* The first group of dependent variables refers to the correctness of the 45 responses per speed condition, which can be expressed either as the average number of correctly recalled targets ( $n$  correct) or as the percentage of trials in which all four targets could be successfully recalled (% correct). Only successful trials in the latter sense were included in the aggregation of gaze-related variables, as in cases when just a subset of targets was correctly identified, participants have probably lost one of the targets somewhere over the trial such that at the point of time the calculation of the virtual centroid would have lost its validity. Thus, the above-defined variable of gaze-overlap time was calculated for trials with perfect target recall only (% of total time). Since no specifics of a certain target or a certain distractor were to be expected, gaze-overlap times were calculated as the averages of one centroid-related, four target-related, and six distractor-related values, respectively.

All dependent variables were analyzed with repeated-measures ANOVAs with the three speed conditions as the within-subject factor, in the case of the gaze-allocation analysis completed by a second within-subject factor for the three regions of interest (centroid, target or distractor). As the only reason for including three motion-onset conditions was to prevent participants from learning of the otherwise to-be-anticipated onset-time of the respective event and as no theory-related predictions can be derived for this variable, the stationary-centroid phase onset was not included as a factor in the ANOVA calculations. Significant interaction effects were further analyzed with planned t-tests. Based on the results reported by Fehd and Seiffert (2010), the alpha level for tests on differences was a priori set to  $\alpha = .05$ . A posteriori effect sizes were computed as partial eta squares ( $\eta_p^2$ ), and in cases of non-significant tests, the power ( $1-\beta$ ) was calculated. According to Fehd and Seiffert (2010), highest response accuracies were expected for the 6 °/s speed condition, followed by the 9 °/s and the 12 °/s conditions. Furthermore, gaze overlap was expected to be highest for the centroid, followed by the target and the distractor average. Finally, again according to the results ob-



tained by Fehd and Seiffert (2010), the gaze behavior should turn out to be independent of stimulus speed.

## Results

As illustrated in Figure 2, results show significant response accuracy differences for the three speed conditions  $F(2, 26) = 62.7, p < .01, \eta_p^2 = .83$ , with the highest accuracies in the 6 °/s condition, followed by the 9 °/s and 12 °/s conditions, and each speed condition differing in terms of response accuracies from each other (all  $ps < .01$ ). The same data pattern is revealed for the average number of correctly identified targets  $F(2, 26) = 54.8, p < .01, \eta_p^2 = .81$ , all  $ps < .01$ , which means that the slower the objects moved, the more targets were correctly recalled by the participants.

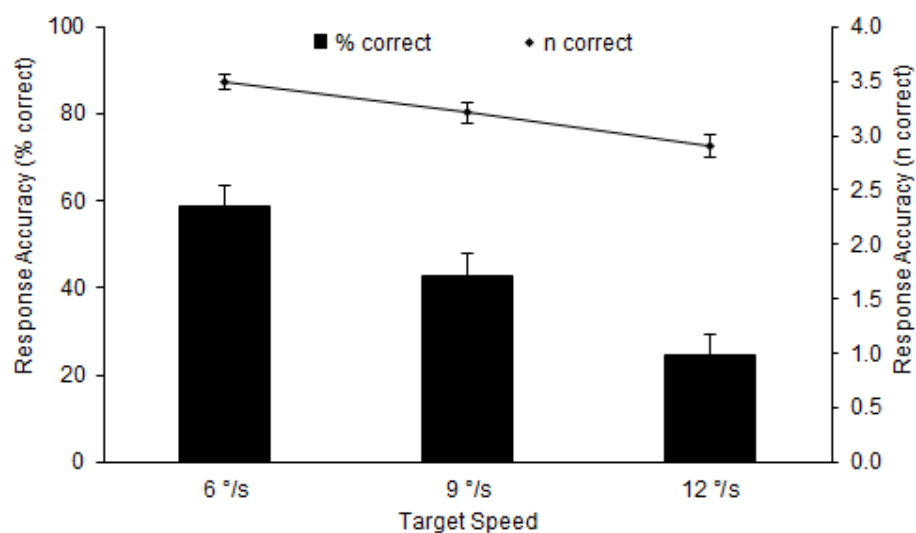


Figure 2. Response accuracy as percentage of correct responses ( $M$  and  $SE$ ) and number of correctly recalled targets ( $M$  and  $SE$ ) as a function of target speed.

In terms of gaze overlap to a priori-defined regions of interest, as depicted in Figure 3, a 3 (speed)  $\times$  3 (region) ANOVA with repeated measures on both factors revealed a significant main effect for region only,  $F(2, 26) = 76.4, p < .01, \eta_p^2 = .86$ . Gaze was allocated more to the centroid than to the targets or the distractors, while gaze overlapped more with targets

than with distractors, all  $ps < .01$ . Neither the effect for speed ( $p = .90$ ,  $1-\beta = .93$ ) nor the interaction of speed and region ( $p = .92$ ,  $1-\beta = .96$ ) reached significance.

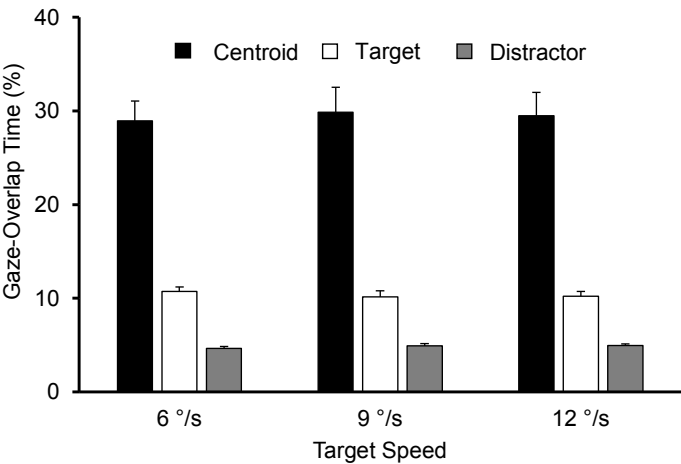


Figure 3. Gaze-overlap time ( $M$  and  $SE$ ) on the centroid, the target average, and the distractor average as a function of target speed.

Discussion

Experiment 1 aimed to answer the question whether classical MOT findings on gaze patterns, in particular those reported by Fehd and Seiffert (2010), could be replicated in consideration of the novel experimental setting with a large projection and the introduction of a stationary-centroid phase. Regarding this replication trial, results show that the faster the objects move, the more tracking accuracy is impaired (see Figure 2). This result is in perfect agreement with previous research (Alvarez & Franconeri, 2007; Fehd & Seiffert, 2010). Compared to the current findings, Fehd and Seiffert (2010) reported higher response accuracies for the conditions 6 °/s (90 %) and 12 °/s (65 %), while the number of correctly tracked targets in our study is comparable with the numbers reported by Alvarez and Franconeri

(2007) at similar speed conditions (i.e., 7 °/s and 14 °/s) when spacing (1 stimulus diameter) is alike (3.4 and 2.4 targets, respectively).

Regarding gaze behavior, Fehd and Seiffert (2010) found that participants were mainly looking at the centroid of the target group, and this was true even at high object speeds. This finding could also be replicated in the current study since results show about 29.5 % gaze overlap with the centroid (Fehd & Seiffert, 2010: 25-30 %), 10.5 % with the target average (Fehd & Seiffert, 2010: approx. 10 %), and 4.8 % with the distractor average (Fehd & Seiffert, 2010: approx. 5 %; see Figure 1B in Fehd & Seiffert, 2010). A further analysis of gaze switches revealed that center-target switches ( $M = 5.5$ ,  $SD = 1.9$ ) were more often used than target-target switches ( $M = 1.3$ ,  $SD = 0.6$ ) which again is in perfect line with previous results (Fehd & Seiffert, 2010: Exp 1: 5.9 vs. 1.9). Thus, it could be demonstrated again that the centroid seems to have a pronounced value in MOT. From a vision perspective, an obvious explanation for this gaze behavior is that peripheral information about target positions are used to effectively monitor all targets. On the basis of this finding, the central research question on the role of peripheral vision in monitoring moving objects regarding unexpected changes can be reasonably asked.

## Experiment 2

In Experiment 2, the functionality of peripheral vision in tracking multiple objects will be under investigation. In particular, the study aims to answer the question whether participants are able to respond to a target-related change in MOT peripherally, that means, before gaze is on this target. Furthermore, since motion information are used for tracking and because of a high temporal resolution of the peripheral retina, it is also predicted that motion changes are detected better than form changes if both kinds of change detections feature a

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

comparable degree of task difficulty. For analyzing events such as motion or form changes, the evaluated stationary-centroid phase from Experiment 1 will be important, as this manipulation opens the door for precisely controlling eccentricities of events. For evaluating the dynamics of the visual search behavior that underlies change identification, gaze will be analyzed using a saccade-detection algorithm to determine the “saccadic reaction time”, which is defined as the time interval between the onset of a change and the onset of the saccade. It is expected to find differences in saccadic reaction times between the two manipulation conditions because if both kinds of change detections are of comparable task difficulty, due to motion sensitivity of peripheral vision, a motion change should be better detected with peripheral vision than form change. More specifically, as peripheral vision is often used to guide saccades (Knowler, 2011) and a rescue saccade is more likely to be used in the form-change condition to perceive the change with overt attention (i.e., while the square is still a diamond), shorter saccadic reaction times are expected in this case, which would also indicate that participants rely less on peripheral vision in the form-change condition.

In Experiment 2, the motion change was implemented as a sudden target stop from a speed of 6 °/s and the form change as a square-to-diamond change over the stationary-centroid phase. These specifications were foremost made for reasons of ecological validity, which means, in reference to change-detection demands in natural settings. When illustrating this point by the sports examples sketched in the beginning of this paper, in a tennis match with an opponent moving close to the court’s opposite baseline, this motion change would correspond to an opponent’s sudden stop of a – rather slow – running movement at a speed of about 2.5 m/s whereas the form change would approximately correspond to an opponent’s racket movement of 30 cm to the left or right side of his or her body. Hence, from an applied perspective, the experimental conditions can definitely claim ecological validity which is im-

portant when it comes to the derivation of recommendations for real-life practice like sports training.

However, based on these specifications and with regard to the above-made predictions on superior detections rates for motion changes, the only point that could be inferred from an empirical confirmation would be that motion changes *as they appear in natural settings* can be better detected than form changes *as they appear in natural settings*. From an experimental point of view, this inference satisfies to a limited degree only, as the alternative explanation that superior motion-change detection rates are mainly caused by higher salience cannot be ruled out. The crucial argument would then be that the specified motion change simply features a minor task difficulty which could be turned around either by a difficulty increase of the motion-change detection task (e.g., slowing down the target speed from 6 °/s to 5 °/s) or by a difficulty decrease of the form-change detection task (e.g., if the target becomes a cross of twice the original size). At this point, an objective measure of the respective task difficulty would be desirable in order to be able to choose motion and form changes of a comparable difficulty. However, due to the dimensional character of the two types of changes, such a simple measure is not available for a principal reason explained by Rensink (2002) in his review on change-detection research as follows: “Performance depends on the magnitude of the change and there is no simple way to equate the visibility of different kinds of changes” (p. 255).

For this reason, in order to empirically control for task-difficulty differences between the two change conditions, additional data was gathered by putting the two types of changes into foveal vision and collecting response times (as detection rates are generally 100 % for foveal vision). Response times of the same magnitude could then be taken as a hint that the specified changes do *not* differ in task difficulty per se such that a potentially revealed superiority of motion-change detection rates by peripheral vision should be foremost ascribed to

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

specifics of the peripheral visual system. In the following, details of this task-difficulty check (TDC) will be reported at the end of each Methods subsection as well as at the beginning of the Results section.

**Methods**

*Participants.* Since performing a MOT task and additionally detecting target changes is a quite complex experimental task, it seemed advisable to conduct the study with participants who were familiar with the basic task of MOT. For this reason, Experiment 2 was run with the participants who had already taken part in Experiment 1 ( $N = 14$ ), accounting the 135 trials of Experiment 1 as familiarization phase for the increasing task demands. Experiment 2 was conducted five weeks after the finalization of Experiment 1. For the additional TDC study, 36 volunteers were recruited (12 females and 24 males; age:  $29.25 \pm 6.65$  years). Both experiments were undertaken in accordance with the Declaration of Helsinki.

*Stimuli.* Stimuli were created in the same way as in Experiment 1 but differed regarding speed, which was set to 6 °/s for all trials since response accuracy at this speed was highest in Experiment 1. Furthermore, motion and form changes of one of the targets were introduced. These changes were evoked in constant relation to the onset of the stationary-centroid phase, which, as in Experiment 1, started at either 3.0 s, 3.5 s, or 4.0 s after motion onset of the objects. Unknown to the participants, the change always occurred exactly 250 ms after the collision of the target with the distractor, which caused the stationary-centroid phase. At this point of time, depending on the trial's specific manipulation condition, one out of the four targets either stopped for 0.5 s before continuing the motion (motion change) or abruptly became a diamond (i.e., the square was rotated by 45 °) for 0.5 s but without altering the current speed (form change) (see Figure 4). Furthermore, to ensure a pre-defined eccentricity of

events, the distance of the event-target was forced to be  $15^\circ$  (if participants, as expectable from Experiment 1, actually pursued a centroid strategy). To prevent participants from anticipation of the changes, a control condition without any target changes was included. As each control trial matched a change trial with a particular motion onset, despite the absence of any change, the control trials were assigned to a (virtual) stationary-centroid phase onset condition in accordance to this match.

The same stimulus configuration was used in the TDC experiment. This time, however, at the beginning of a single trial, only one target was highlighted with a red frame, namely the target which would (virtually) change its motion or form over the subsequent motion phase. Participants were instructed to keep this target in foveal vision during the whole trial.

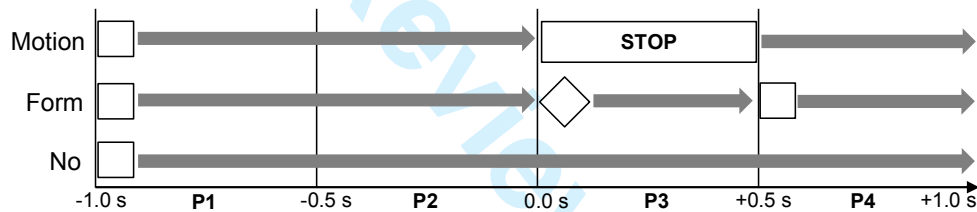


Figure 4. Temporal properties of the three target-change conditions. In phase P3, the target either stopped (Motion) or became a diamond (Form) or continued to move unchanged (No) for 0.5 s whereas no differences between the conditions are evident for the remaining phases P1, P2, and P4.

*Apparatus.* In addition to the setup used in Experiment 1, a single-button response device (1000 Hz) was integrated into the VICON data collection system to measure motor response times.

*Procedure.* Experiment 2 was organized in individual single sessions that lasted about one hour. Participants' task was to press the button as soon as a motion- or form-related target change occurred and to name the number of the changed target at the end of the trial. However, if no change occurred, the task was to recall all four targets. Participants initially practiced each of the three conditions with three trials that were not shown later in the test blocks, re-

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

sulting in 9 practice trials in total. Taking the stimulus constraints into account, a total of 120 trials were derived from 12 mother trials. These trials were quasi-randomly ordered in 10 blocks, with 12 trials each with the additional specification that per block each mother trial appeared exactly once and each change/control condition and each onset of the stationary-centroid phase exactly four times. As in Experiment 1, verbal decisions were recorded in writing, and no knowledge of the results was given after the test trials.

In the TDC experiment, participants started with 6 familiarization trials followed by 30 test trials in 2 blocks (10 trials for form, motion and no-change condition each) in which they had to press the button as soon as a motion- or form-related target change occurred. In addition, at the end of each trial, participants had to name the number of the target they had followed with their eyes.

*Gaze analysis.* Since in Experiment 2 the focus was laid on the dynamics of gaze behavior as a function of monitoring object changes with peripheral vision, two kinds of trials had to be excluded from further analysis. First, a trial could not be included if the target change was not correctly signaled by a button press or not correctly named at the end of the trial's motion phase, as in those trials attention was not on the correct target. Second, trials were excluded from further analysis if, at event start, the distance of the gaze vector to the to-be-changed target was less than 5 °, that is, within the range of parafoveal vision (Calvo & Lang, 2005). The reason for this exclusion is that in those trials the participants could have focused on the respective target just by chance so that the missing need for a saccade should be considered as accidental and not be interpreted as resulting from a certain gaze-control strategy. In the Results section, detailed information will be given on the percentage of trials that had been excluded due to these criteria.



For the remaining trials, two gaze-related groups of variables were aggregated. First, the distance of the gaze vector to the target as well as to the centroid was calculated (in  $^{\circ}$ ) as the average value for four phases P1-P4 of 0.5 s duration each. As depicted in Figure 4, these phases were defined in relation to the change event (in case of control trials, to the virtual change event) with P1 beginning 1.0 s before, P2 beginning 0.5 s before, P3 beginning exactly with the target change and P4 beginning exactly with the change's termination. After the calculation of the mean gaze-target and gaze-centroid distances for the four phases per single trial, the values were averaged for each participant over the valid trials per change/control condition.

To analyze the dynamics of the peripheral monitoring of the targets, for the valid motion- and form-change trials, a second group of gaze-related variables referred to the onset and offset of the first saccade to the changed target. For this purpose, saccades were identified by a velocity-based detection algorithm with adaptive thresholds on the basis of local noise levels (Nyström & Holmqvist, 2010). This analysis resulted in three time intervals (in ms), which (in cases in which the identification was signaled by the participants after the saccade) accumulated to the motor-response time indicated by the button press: a saccadic reaction time from the onset of the target change to the onset of the saccade, a saccadic duration from the onset to the offset of the saccade, and a fixation duration from the offset of the saccade to the button press. For each participant, the values per trial were averaged over the valid trials for motion- and form-related changes. It can be argued that the shorter the saccadic reaction is, the less is the change detection based on peripheral vision.

In the TDC experiment, it was controlled that participants used foveal vision for detecting target changes (maximum distance of  $3^{\circ}$  of visual angle to the target).

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

*Measures.* Besides the gaze-related variables, response accuracy (in %) was gathered, in the no-change trials as the percentage of trials in which all four targets could be correctly recalled, and in the motion- and form-change trials as the percentage of trials in which the target change was correctly signaled by the button press and the correct number of the event-target was named at the end of the trial as well. Furthermore, motor-response time was calculated (in ms) for all correct motion- and form-change trials as the time between target-change onset and button-press and then averaged over all available trials. Response time calculations in the TDC experiment were conducted in the same manner.

All dependent variables were analyzed with repeated-measures ANOVAs, including either all three or just the two target-change conditions as a first factor and, if applicable, a gaze- or phase-related variable as a second or third factor. Significant main or interaction effects were further analyzed with paired t-tests. The alpha level for tests on differences was a priori set to  $\alpha = .05$ , a posteriori effect sizes were computed as partial eta squares ( $\eta_p^2$ ), and the power ( $1-\beta$ ) was calculated in cases of non-significant tests.

It was expected to find higher response accuracies in the motion- compared to the form-change condition, while response accuracies in the no-change condition should not differ from the results obtained in Experiment 1 for the same stimulus speed. Furthermore, it was hypothesized that the average gaze distance to the centroid is smaller than to the manipulated target in P1 and P2. However, gaze should be closer to the target in the form- and motion-change condition in P4 (latest), whereas in the no-change condition, the gaze distance should be continuously smaller to the centroid than to the target. For the TDC experiment, it was predicted that the participants were *not* able to respond faster to foveally perceived motion changes than to foveally perceived form changes. Under the assumption that this prediction was confirmed, a superior performance for motion than for form changes was anticipated for the peripheral viewing conditions of the main experiment. With respect to the dependent vari-

ables described above, it was particularly expected that the motor response can be executed ahead of the saccade onset in more trials for the motion-change than in the form-change condition.

587

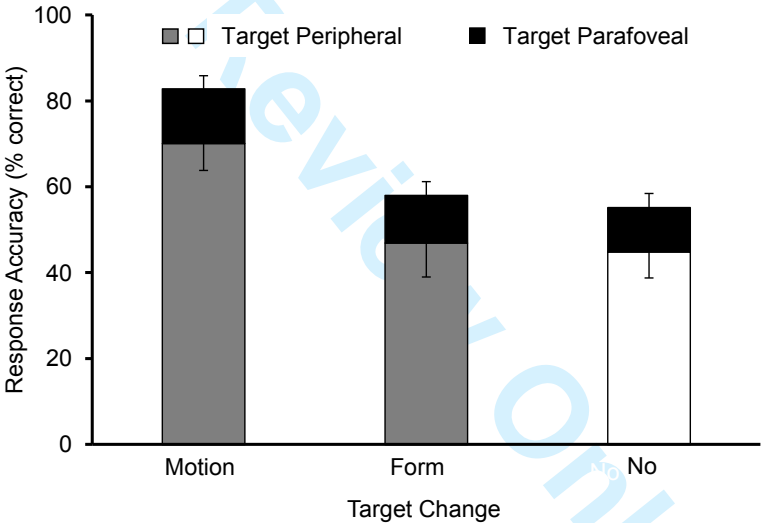
## 588 Results

589 For the TDC experiment, paired t-tests were used to test response time differences for  
590 motion vs. form change detections and Cohen's  $d$  was calculated as effect size. Data of 35  
591 participants could be analyzed. Results not only revealed that motion changes were *not* de-  
592 tected faster than for form changes, as it has been predicted above, but that form changes were  
593 actually detected significantly faster than motion changes when being put in foveal vision  
594 (motion:  $M = 427.8$  ms,  $SE = 72.3$  ms; form:  $M = 289.1$  ms,  $SE = 48.9$  ms),  $t(34) = 11.39$ ,  $p <$   
595  $.01$ ,  $d = 1.7$ .

596 For the main experiment, three participants had to be excluded from the analysis be-  
597 cause they were only able to perceive a target change if central gaze was by chance on the  
598 respective target leading to less than five trials that fulfilled the above-defined inclusion crite-  
599 ria. On the contrary, the remaining 11 participants clearly pursued a centroid-tracking strate-  
600 gy, as at target-change onset, the respective target was out of parafoveal vision in 83.2 % ( $SE$   
601  $= 4.1$  %) of the correctly identified motion-change and in 78.0 % ( $SE = 6.4$  %) of the correctly  
602 identified form-change trials, respectively. When applying the same calculation to the corre-  
603 sponding no-change conditions, the target that would change in the change-derivatives of the  
604 same mother trial, one receives an average of 79.4 % ( $SE = 6.4$  %) of trials in which all the  
605 four targets could be successfully recalled and the respective target was out of parafoveal vi-  
606 sion at the moment of the (virtual) target change. To provide consistency, the following re-  
607 sults refer to the 11 selected participants and to the just-described percentage of valid trials.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

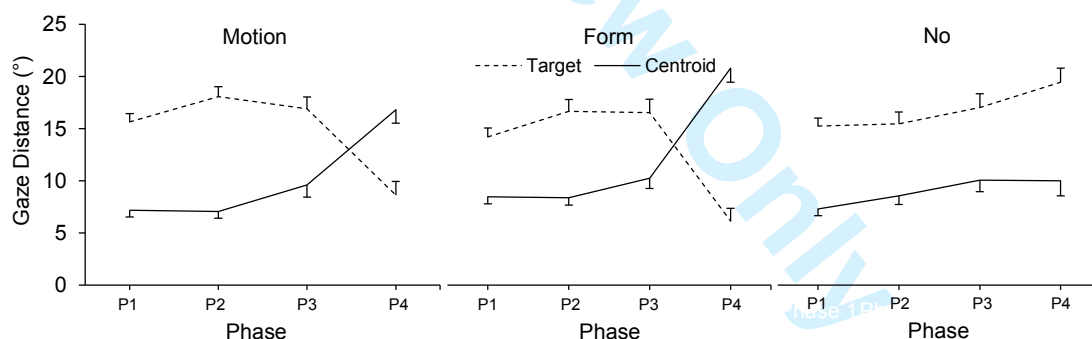
608 As depicted in Figure 5, participants were able to detect target-motion changes better  
609 than form changes. This is confirmed when the ANOVA is based on all trials in which the  
610 change could be identified correctly (grey plus black fractions of the bars),  $F(1, 10) = 17.2, p$   
611  $< .01, \eta_p^2 = .63$ , as well as when the calculation is restricted to the trials that were valid in that  
612 sense that, at target-change onset, the respective target was out of parafoveal vision (grey  
613 fractions of the bars),  $F(1, 10) = 10.71, p < .01, \eta_p^2 = .52$ . Furthermore, response accuracy in  
614 the no-change condition ( $M = 55.2 \%$ ,  $SE = 4.5 \%$ ; white plus black fraction of the bar)  
615 showed no significant differences to those obtained in Experiment 1 for the same object speed  
616 ( $M = 59.6 \%$ ,  $SE = 6.1 \%$ ),  $t(11) = 0.82, p = .43, 1-\beta = .74$ .



617  
618 *Figure 5.* Response accuracy ( $M$  and  $SE$ ) as a function of target-change condition. In the motion- and form-  
619 change conditions, a trial was counted as correct if the button was pressed and the correct target number was  
620 named at the end of the trial whereas, in the no-change condition, all four targets had to be recalled correctly.  
621 The black fractions of the bars depict the percentage of trials that were excluded from further analyses because,  
622 at change onset, the respective target was already in parafoveal vision.

623 For the event-related gaze-distance analysis, which is illustrated in Figure 6, a 3 (target  
624 change) x 2 (location) x 4 (phase) ANOVA with repeated measures on all three factors was  
625 conducted. The results show significant main effects for location,  $F(1, 10) = 8.21, p = .02, \eta_p^2$

626 = .45, and phase,  $F(3, 30) = 30.7$ ,  $p < .01$ ,  $\eta_p^2 = .76$ , as well as the following significant two-  
 627 way interactions: target change x location,  $F(2, 20) = 21.3$ ,  $p < .01$ ,  $\eta_p^2 = .68$ , target change x  
 628 phase,  $F(6, 60) = 4.9$ ,  $p < .01$ ,  $\eta_p^2 = .33$ , and location x phase,  $F(3, 30) = 42.6$ ,  $p < .01$ ,  $\eta_p^2 =$   
 629 .81. Furthermore, a significant three-way interaction target change x location x phase was  
 630 found,  $F(6, 60) = 83.8$ ,  $p < .01$ ,  $\eta_p^2 = .89$ . Consequently, to further interpret the three-way  
 631 interaction, paired t-tests were used to evaluate gaze-distance differences to the centroid com-  
 632 pared to the manipulated target for all target-change conditions and phases. The results show  
 633 that, for all target-change conditions, the gaze vector was closer to the centroid than to the  
 634 respective target in the first three phases (all  $ps \leq .01$ ), while in P4, gaze was closer to the  
 635 target in the motion-change condition as well as in the form-change condition (all  $ps < .01$ ),  
 636 whereas the opposite was true in the no-change condition ( $p < .01$ ). Furthermore, the gaze  
 637 vector was closer to the target in P4 of the form trials compared with P4 of the motion-change  
 638 trials ( $p = .03$ ).



639  
 640 *Figure 6.* Gaze distance ( $M$  and  $SE$ ) to the centroid and the changing target for four phases (P1-P4), separately  
 641 depicted for three target-change conditions. In the no-change condition, gaze distances were calculated to the  
 642 target that changed in the matched trials of the other conditions.

643 In Figure 7, the dynamics of the interplay between peripheral monitoring and saccadic  
 644 reaction from the centroid to the changed target is illustrated in more details. As depicted in  
 645 Figure 7A, the button was pressed before gaze was on the target in a higher percentage of

correct trials in the motion-change condition compared to the form-change condition,  $F(1, 10) = 8.9, p = .01, \eta_p^2 = .47$ , which means that in those cases, the change was definitely detected by means of peripheral vision. Furthermore, as shown in Figure 7B, a respective ANOVA reveals significant differences in saccadic reaction time after target-change onset,  $F(1, 10) = 6.7, p = .03, \eta_p^2 = .40$ , as well as in fixation time before the motor response was indicated by the button press,  $F(1, 10) = 7.1, p = .02, \eta_p^2 = .42$ , whereas no differences were found for saccade duration,  $F(1, 10) = 1.6, p = .24, 1-\beta = .94$ . This means that later saccade onsets to the changed target were observed in the motion-change condition compared to the form-change condition, while in the form-change condition, targets were fixated longer before the motor response was executed.

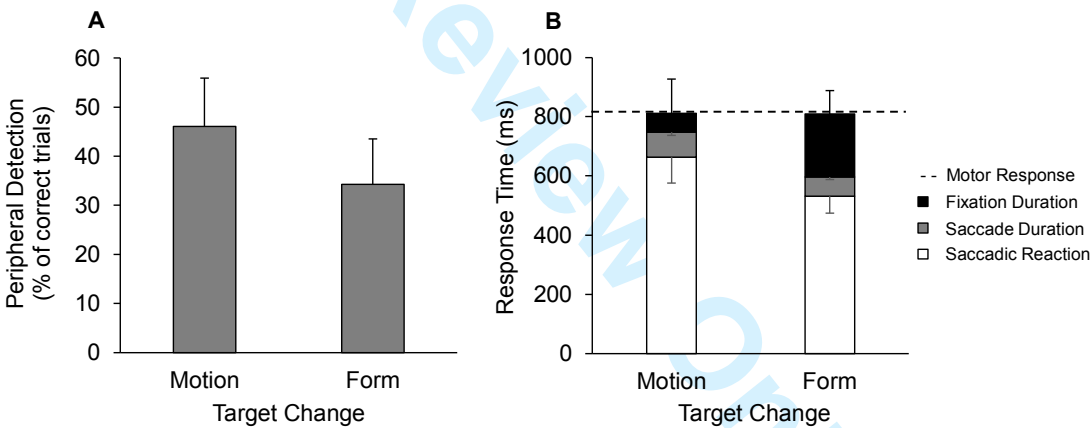


Figure 7. (A) Peripheral change detection ( $M$  and  $SE$ ) as a function of target change, calculated as the percentage of trials in which the motor response was executed before the gaze vector reached the respective target. (B) Saccadic reaction, saccade duration, and fixation duration ( $M$  and  $SE$ ) as cumulated response up to the average motor response time. The value “0” on the vertical axis signifies the onset of the target change.

## Discussion

In Experiment 2, a novel approach in studying MOT performance has been introduced by using an event-driven saccade-based gaze analysis that precisely allows for examining

whether motor responses can be executed when changes in the environment are perceived with peripheral vision only. First of all, for trials without a change, intra-individual response accuracy did not differ from the values obtained in Experiment 1, which could be seen as an indicator that the additional detection task did not affect tracking performance negatively (as it was also suggested by Bahrami, 2003). Furthermore, no learning effects were observed over the five weeks between both experiments. On this basis, with respect to the main research question, it could be clearly shown that motor responses can be reliably initiated if the decisive stimuli can be perceived by peripheral vision only. Moreover, as hypothesized, higher response accuracies for the motion-change condition could be revealed than for the form-change condition. For the appropriate interpretation of this finding the results of the TDC study must be taken into account as they show faster responses to form than to motion changes (of exactly the same magnitude as in the main experiment) when the changes are put into foveal vision such that the superior detection of motion changes in the main experiment can definitely not be ascribed to a mere task-difficulty effect but must be attributed to specifics of the peripheral visual system.

Having a closer look onto the main results, the gaze distance was significantly closer to the centroid than to the targets over the two phases before the event (P1 and P2) and even over the phase of the change (P3), whereas over the last phase (P4), gaze was mostly on the target in change trials whereas it stayed closer to the centroid in no-change trials. At this point, it should be noted that the gaze distances depicted in Figure 6 resulted from an averaging procedure which means that a relatively small percentage of trials with no saccades to the target result in large gaze-target distances in P4 of the change conditions. The same is true for the gaze-centroid distances in P1-P3 and, in addition, in P4 of the no-change condition since saccades to targets are occasionally initiated (Elfanagely et al., 2011; Fehd & Seiffert, 2008, 2010), thereby increasing the mean distance to the centroid. For this reason, the respective

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

values ranging between 5 ° and 10 ° should, in the end, be understood as a result of focusing either a target or the centroid in most of the cases, but not always. Hence, the data should be interpreted in such a way that, in cases of a correct response, participants produce a saccade from the centroid to the respective target whenever a target changes its motion or form.

However, when bearing in mind that the target change over P3 lasted 0.5 s only and that the change was terminated at the beginning of P4, this finding should not be misinterpreted in such a way that the saccades revealed for P4 are mandatorily needed to detect the respective change. This conclusion is underpinned by the gaze data depicted in Figure 7, showing that mean saccadic-reaction times are longer than 0.5 s in both target-change conditions and that, in a considerable percentage of trials, the saccade was initiated even after the button press, that is, on average, more than 0.8 s after the target change. Hence, in a lot of cases, the decision must have been made before a verification by a suitable fixation could have taken place.

When comparing the two target-change conditions in more details, in line with our predictions, in significantly more correct trials in the motion-change compared with the form-change condition, the button press had been executed before the saccade to the changed target was initiated. In both cases, as already argued above, these changes must have been perceived with peripheral vision since the gaze position was more than 5 ° away from the changing target. Furthermore, saccadic reaction time turned out to be longer in the motion-change compared with the form-change condition supporting the assumption that foveal information are not as much required for motion-change as for form-change detection. On the contrary, as hypothesized, a time of gaze being on the changed target seems to be required in the form-change condition before the button press can be executed (which would also be in line with the results of the TDC experiment that form changes can be detected faster with foveal vision).



To test whether this saccade is helpful to respond faster, a “saccading strategy” (i.e., gaze is on the target *before* the button is pressed) and a “peripheral strategy” (i.e., gaze is on the target *after* the button is pressed) can be distinguished and checked against each other by post-hoc paired t-tests for both manipulation conditions and gaze strategies. Eight participants could be integrated into this analysis because they exhibited correct responses with both gaze strategies in both conditions. Results show that response times for the two gaze strategies neither differ in the form- ( $p = .93$ ) nor in the motion-change condition ( $p = .40$ ), indicating that initiating a saccade to the target does *not* help to respond faster compared to using peripheral vision only. Thus, initiating a saccade to the target does *not* lead to faster responses but could be interpreted as just being helpful to early track the target for naming the correct number at the end of the trial.

## Overall Discussion

The current study aimed to investigate whether peripheral vision can be used to monitor multiple moving objects as well as to detect single-target changes. For this purpose, the designated experimental MOT setup with a large projection and a constant-position centroid phase had to be checked first. Therefore, in Experiment 1, the successful replication of the findings regarding the use of a virtual centroid to track multiple objects (c.f., Fehd & Seiffert, 2008, 2010) and the dependency of tracking accuracy on target speed (c.f., Cavanagh & Alvarez, 2005) had been important for conducting the main study. In the subsequent Experiment 2, it could be clearly shown that peripheral vision is used at great extents to detect changes in MOT. Moreover, for the change conditions specified on the basis of considerations regarding the ecological validity of the manipulations, it could be demonstrated that motion changes can be better detected peripherally than form changes. Finally, an additionally conducted control

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

experiment showed that this superiority must be explained by particularities of the peripheral visual system and cannot be attributed to possible differences in task difficulty.

As debated in the introductory section, findings on MOT are often discussed exclusively from an attention perspective. However, based on the gaze-related findings revealed in the experiments at hand, it seems worthwhile to consider classical MOT results also from a vision perspective. In this regard, peripheral vision (i.e., perceiving stimuli with peripheral vision) and spatial covert attention (i.e., allocating attention without eye movements; Strasburger et al., 2011) seem to describe closely related phenomena. Therefore, in the following paragraphs, five reasons will be discussed regarding an integration of properties of the visual system in the context of attention-related findings to understand the role of covert attention in MOT. These reasons refer to (a) the ecological validity of methods to measure covert attention, (b) the detection of feature changes that depend on visual capabilities, (c) the connection of covert attention to a centroid-MOT strategy and consequences of saccade-related costs for processes of attentional monitoring, (d) the fact that eye-tracking devices can be generally useful for testing assumptions made in attentional theories on MOT, and (e) spatial uncertainty in the context of covert-attention processes, which could also be related to the voluntary control of saccades. At the end of each paragraph, a follow-up question for future experiments will be derived.

(a) In classical MOT studies, participants were usually instructed to fixate a central point to examine covert attention in detecting feature changes (e.g., Sears & Pylyshyn, 2000). This approach, however, seems problematic since, first, the fixation instruction was not checked by the application of gaze-registration methods and, second, fixating a central point may not reflect the participant's natural behavior. In contrast, in the current experiment, participants were tested under ecologically valid conditions, being free to move their eyes and their head because the actual gaze could be controlled by eye- and head-tracking devices. Be-

764 yond these advantages regarding ecological validity, the event-driven saccade-based gaze  
765 analysis offers the opportunity to estimate whether peripheral or foveal vision was used to  
766 detect changes leading to enriched inferences on the role of covert attention in MOT. Thus, in  
767 future research, by applying eye-tracking methods, covert-attention processes may be exam-  
768 ined even in real-life settings in which free gaze behavior is allowed.

769 (b) The use of covert attention not only depends on psychological but also on visual  
770 capabilities, especially when target changes have to be detected in MOT. In former studies  
771 (e.g., Bahrami, 2003) participants regularly showed better detection rates for color changes  
772 compared with form changes, but, as a matter of the respectively pursued research questions,  
773 visual capabilities were not taken into account. Thus, it is not known so far whether the supe-  
774 riority regarding color-change detection actually results from specifics of attentional process  
775 or could be attributed to particularities of the visual system. In the current study, those visual  
776 capabilities were taken into account because finding better detection rates of motion com-  
777 pared with form changes was expected due to the light and motion sensitivity of the peripher-  
778 al human retina, which comes along with a higher sensitivity to detect motion changes (Fin-  
779 lay, 1982; Goldstein, 2010; McKee & Nakayama, 1984). Indeed, the frequency of detection  
780 was found to be higher in the motion-change condition; however, the contrast sensitivity of  
781 the peripheral retina obviously allowed detecting form changes, too. Consequently, the lower  
782 detection rates for form changes can be traced back to the lower spatial resolution of the pe-  
783 ripheral visual system. Hence, in a follow-up experiment, the low spatial acuity but high mo-  
784 tion sensitivity of the peripheral retina should be further investigated by manipulating the ec-  
785 centricities of changes. In the case of a visual-system effect on detection rates, the prediction  
786 can be derived that, as a function of eccentricity, form-change detection should be impaired to  
787 a higher degree than motion-change detection.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

(c) Previous research has shown that covert attention is used for tracking purposes (Fehd & Seiffert, 2008, 2010). The gaze-related findings of the current study extend this view by showing that a centroid strategy is preferred not only for tracking, but also for change detection. This gaze strategy is presumably favored because a target-switching strategy with an alternating gaze between targets (Fehd & Seiffert, 2010) would cause a lot of costs due to the suppression of information processing before, during, and after the saccade (Diamond, Ross, & Morrone, 2000). As a consequence, the chances to miss the target-feature change would be higher. Participants in our study indeed seem to avoid costly eye movements by using the virtual centroid strategy, in attentional terms: by monitoring the targets with covert attention and “waiting” for a feature change. With the results at hand it can be shown that covert attention is not only used to monitor the distances between targets and barrier occluders (Zelinsky & Todor, 2010) but also to detect target changes. In future experiments, this hypothetical mechanism could be investigated further by manipulating the costs of eye movements. In more details, due to the suppression of information processing as a consequence of saccades, it would be expected that the more saccades are experimentally induced, the more likely a change will remain undetected.

(d) By studying natural gaze behavior, assumptions made by theoretical models on the role of attention in MOT can be assessed more rigidly. On the one hand, our findings confirm attentional theories in different respects. In this regard, the observed gaze behavior indicates the value of the center of the polygon for MOT performance (Fehd & Seiffert, 2010), which is in line with the model proposed by Yantis (1992). Furthermore, the higher detection rates in the motion-change condition can be explained by the fact that the anticipated target locations of the polygon (Iordanescu, Grabowecky, & Suzuki, 2009) are no longer perceptible because of a target stop, whereas the polygon still exists in the form-change condition causing lower detection rates. In this case, presumably because of the high tracking load with four targets

(Fencsik et al., 2007, Howe & Holcombe, 2012; Luu & Howe, 2015), no motion extrapolation seems to be used in MOT. Hence, change detection results could be better explained with the multifocal theory and concurrent streams on the targets where position information are used for tracking as participants were able to detect motion as well as form changes. Thus, a position tracker and a stream of object features seem to be concurrently used to keep track of the targets. On the other hand, however, our gaze analyses also pose challenges for attentional MOT theories. For example, in the model proposed by Cavanagh and Alvarez (2005), the relative impact of the motion tracker and feature stream has to be questioned, as it can be inferred from the results at hand that, when using covert attention to keep track of the targets, the position tracker has a more pronounced sensitivity in the concurrent-tracking procedure because motion changes were detected more often than form changes. This line of thought would be supported by the weighting-mechanism proposed by Papenmeier et al. (2014) because spatiotemporal information during tracking are reliable until a motion change is induced but a “false” spatiotemporal reliability is given in the form change condition. Hence, the decrease in distinctiveness of spatiotemporal and feature information could explain lower detection rates in the form change condition. Besides the relation of both streams, the concurrency of attentional streams was challenged with our results because a number of changes were not detected. Thus, it could be that streams are not permanently existent. Instead, covert attention seems to switch between the targets leading to a limited amount of time to process target information which, in turn, would increase the difficulty to detect a change and explain missed changes. Another challenge arises in the context of the FINST-model, in which it is assumed that focal attention must be directed to the targets to update their positions (Fehd & Seiffert, 2010; Oksama & Hyönä, 2004). Contrary to this assumption, our research shows that peripheral vision, that is, covert attention, can be used to update target features and to detect feature changes whereas foveal vision, that is, focal attention, seems mainly to be used to keep track

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

838 of the changed target to correctly name its number at the end of the trial. In particular, the  
839 assumption of the feature-blind visual index model by Pylyshyn (1989, 2007), that feature  
840 information cannot be accessed during tracking, is challenged by our data. On the contrary,  
841 our results support the findings by Papenmeier et al. (2014), that spatiotemporal *and* feature  
842 information can be processed. Consequently, future research on MOT should be directed to  
843 the empirical test of specific predictions of attention-related models by integrating eye-  
844 tracking technology. In particular, it should be tested whether it is the stop of the anticipated  
845 target location that causes its detection or whether a slowdown would lead to the same effect  
846 to figure out how closely the polygon is linked to the mechanisms of attentional tracking.

(e) Our findings on gaze behavior in MOT also give rise to the speculation that the use  
of overt attention, measured as saccadic reaction times, is affected by spatial uncertainty as  
well as by the voluntariness of saccades. First, the ability to use covert attention to detect  
changes in the environment can be related to spatial uncertainty because a low contrast in pe-  
ripheral stimuli results in higher spatial uncertainty (for an overview, see Carrasco, 2011). In  
the current experiment, form changes presumably caused less contrast changes per frame so  
that spatial uncertainty can be assumed to be higher in this condition. As a consequence, the  
early saccade onset in the form-change condition might result from this uncertainty. In con-  
trast, in the motion-change condition, the saccade was initiated later, potentially because of  
lower spatial uncertainty. Thus, it seems worthwhile to take spatial uncertainty into account in  
future experiments on covert attention in MOT. Furthermore, it can be speculated that sac-  
cades were voluntarily controlled. This speculation is substantiated by the fact that saccades  
are usually initiated in a bottom-up process resulting in an involuntary overt-attentional shift  
as soon as stimuli in the environment change their motion (Drew et al., 2009; Hillstrom &  
Yantis, 1994). In our case, those shifts of overt attention were only sparsely observed. Hence,  
the assumption seems to make sense that the saccades to the event-target were voluntarily

controlled (Peterson, Kramer, & Irwin, 2004) in order to update the target position because of the higher foveal resolution (Zelinsky & Todor, 2010). Consequently, in a follow-up study, it is planned to eliminate the task-dependent evocation of a saccade by pursuing a dual task approach in which a button has to be pressed as soon as a change is detected and, regardless of a change, all four targets have to be recalled at the end of the trial. If peripheral vision sufficiently allows for change detection, it can be predicted that no or at least less saccades will be executed because the main task would be to identify the four targets at the end of each trial. If, by this means, the independence of change detection from subsequent saccades could be confirmed, it is also planned to experimentally vary the before discussed uncertainty of the change event to investigate whether the degree of uncertainty plays a crucial role in the effectiveness of peripheral monitoring in MOT.

In summary, our experimental approach allows for a more ecologically valid examination on covert attention in MOT. Our empirical results show that the detection of feature changes depend on visual capabilities like spatial resolution and motion sensitivity, that the centroid-strategy for MOT seems also to be beneficial for change detection to avoid saccade-related costs, that feature and spatiotemporal information can be processed with covert attention by concurrent attentional streams, and that spatial uncertainty might affect the voluntary control of change detection with either overt or covert attention. In any case, the integration of eye-tracking methods in the field of attention-related MOT research seems to be worthwhile as existing findings may appear in a new light and novel research questions may reasonably be derived. Hence, follow-up questions that have been suggested above are currently investigated in our research group.

885



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**References**

Alvarez, G. A., & Franconeri, S. L. (2007). How many objects can you track? Evidence for a resource-limited attentive tracking mechanism. *Journal of Vision*, 7, 1-10.

Bahrami, B. (2003). Object property encoding and change blindness in multiple object tracking. *Visual Cognition*, 10, 949-963.

Boucart, M., Moroni, C., Thibaut, M., Saffarczyk, S., & Greene, M. (2013). Scene categorization at large visual eccentricities. *Vision Research*, 86, 35-42.

Calvo, M. G., & Lang, P. J. (2005). Parafoveal Semantic Processing of Emotional Visual Scenes. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 502-519.

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, 51, 1484-1525.

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends in Cognitive Sciences*, 9, 349-354.

Catteeuw, P., Helsen, W., Gilis, B., Van Roie, E., & Wagemans, J. (2009). Visual scan patterns and decision-making skills of expert assistant referees in offside situations. *Journal of Sport and Exercise Psychology*, 31, 786-797.

Davids, K. (1984). The role of peripheral vision in ball games: Some theoretical and practical notions. *Physical Education Review*, 7, 26-40.

Diamond, M. R., Ross, J., & Morrone, M. C. (2000). Extraretinal control of saccadic suppression. *Journal of Neuroscience*, 20, 3449-3455.

Drew, T., McCollough, A. W., Horowitz, T. S., & Vogel, E. K. (2009). Attentional enhancement during multiple-object tracking. *Psychonomic Bulletin & Review*, 16, 411-417.



- 908 Elfanagely, O., Haladjian, H., Aks, D., Kourtev, H., & Pylyshyn, Z. (2011). Eye-movement  
909 dynamics of object-tracking. *Journal of Vision*, 11, 1-16.
- 910 Faubert, J., & Sidebottom, L. (2012). Perceptual-Cognitive Training of Athletes. *Journal of*  
911 *Clinical Sport Psychology*, 6, 85-102.
- 912 Fehd, H. M., & Seiffert, A. E. (2008). Eye movements during multiple object tracking: Where  
913 do participants look? *Cognition*, 108, 201-209.
- 914 Fehd, H. M., & Seiffert, A. E. (2010). Looking at the center of the targets helps multiple ob-  
915 ject tracking. *Journal of Vision*, 10, 1-13.
- 916 Fencsik, D. E., Klieger, S. B., & Horowitz, T. S. (2007). The role of location and motion in-  
917 formation in the tracking and recovery of moving objects. *Perception & Psychophysics*,  
918 69, 567-577.
- 919 Finlay, D. (1982). Motion perception in the peripheral visual field. *Perception*, 11, 457-462.
- 920 Franconeri, S. L., Lin, J. Y., Pylyshyn, Z. W., Fisher, B., & Enns, J. T. (2008). Evidence  
921 against a speed limit in multiple-object tracking. *Psychonomic Bulletin & Review*, 15,  
922 802-808.
- 923 Goldstein, E. B. (2010). *Sensation and perception* (8th ed.). Belmont, Calif: Wadsworth Cen-  
924 gage Learning.
- 925 Gralla, V. (2007). *Peripheres Sehen im Sport – Möglichkeiten und Grenzen dargestellt am*  
926 *Beispiel der synchronoptischen Wahrnehmung [Peripheral vision in sports – Possibilities*  
927 *and limits depicted with the example of synchronoptic perception]*. Dissertation, Ruhr  
928 University Bochum. URL: [http://www-brs.ub.ruhr-uni-bochum.de/netahtml/HSS/Diss/](http://www-brs.ub.ruhr-uni-bochum.de/netahtml/HSS/Diss/GrallaVolker/diss.pdf)  
929 [GrallaVolker/diss.pdf](http://www-brs.ub.ruhr-uni-bochum.de/netahtml/HSS/Diss/GrallaVolker/diss.pdf).

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

930 Hartmann, E., Lachenmayr, B., & Brettel, H. (1979). The peripheral critical flicker frequency.  
931 *Vision Research*, 19, 1019-1023.

932 Helsen, W. F., & Starkes, J. L. (1999). A multidimensional approach to skilled perception and  
933 performance in sport. *Applied Cognitive Psychology*, 13, 1-27.

934 Hillstrom & Yantis (1994). Visual motion and attentional capture. *Perception & Psychophys-*  
935 *ics*, 55, 399-411.

936 Howe, P. D., & Holcombe, A. O. (2012). Motion information is sometimes used as an aid to  
937 the visual tracking of objects. *Journal of Vision*, 12, 1–10.

938 Huff, M., & Papenmeier, F. (2013). It is time to integrate: The temporal dynamics of object  
939 motion and texture motion integration in multiple object tracking. *Vision Research*, 76,  
940 25–30.

941 Huff, M., Papenmeier, F., Jahn, G., & Hesse, F. (2010). Eye movements across viewpoint  
942 changes in multiple object tracking. *Visual Cognition*, 18, 1368–1391.

943 Iordanescu, L., Grabowecky, M., & Suzuki, S. (2009). Demand-based dynamic distribution of  
944 attention and monitoring of velocities during multiple-object tracking. *Journal of Vision*,  
945 9, 2-23.

946 Kowler, E. (2011). Eye movements: The past 25 years. *Vision Research*, 51, 1457–1483.

947 Kredel, R., Klostermann, A., Lienhard, O., Koedijker, J., Michel, K., & Hossner, E. J. (2011)  
948 Perceptual Skill Identification in a Complex Sport Setting. *BIO Web of Conferences*, 1,  
949 00051. URL: <http://dx.doi.org/10.1051/bioconf/20110100051>.

950 Landry, S. J., Sheridan, T. B., & Yufik, Y. M. (2001). A methodology for studying cognitive  
951 groupings in a target-tracking task. *IEEE Transactions on Intelligent Transportation Sys-*  
952 *tems*, 2, 92-100.

- Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the absence of attention. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 9596-9601.
- Luu, T., & Howe, P. D. (2015). Extrapolation occurs in multiple object tracking when eye movements are controlled. *Attention, Perception, & Psychophysics*. URL: <http://link.springer.com/article/10.3758%2Fs13414-015-0891-8>.
- Meyerhoff, H. S., Papenmeier, F., & Huff, M. (2013). Object-based integration of motion information during attentive tracking. *Perception*, *42*, 119–121.
- McKee, S. P., & Nakayama, K. (1984). The detection of motion in the peripheral visual field. *Vision Research*, *24*, 25-32.
- Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eye tracking data. *Behavior Research Methods*, *42*, 188-204.
- Oksama, L., & Hyönä, J. (2004). Is multiple object tracking carried out automatically by an early vision mechanism independent of higher-order cognition? An individual difference approach. *Visual Cognition*, *11*, 631-671.
- Papenmeier, F., Meyerhoff, H. S., Jahn, G., & Huff, M. (2014). Tracking by location and features: Object correspondence across spatiotemporal discontinuities during multiple object tracking. *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 159–171.
- Peterson, M. S., Kramer, A. F., & Irwin, D. E. (2004). Covert shifts of attention precede involuntary eye movements. *Perception & Psychophysics*, *66*, 398-405.
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, *3*, 179-197.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

976 Rensink, R. A. (2002) Change detection. *Annual Review of Psychology*, 53, 245-277.

977 Sears, C. R., & Pylyshyn, Z. W. (2000). Multiple object tracking and attentional processing.  
978 *Canadian Journal of Experimental Psychology*, 54, 1-14.

979 St.Clair, R., Huff, M., & Seiffert, A. E. (2010). Conflicting motion information impairs multi-  
980 ple object tracking. *Journal of Vision*, 10, 1–13.

981 Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recogni-  
982 tion: a review. *Journal of Vision*, 11, 1-82.

983 Thorpe, S. J., Gegenfurtner, K. R., Fabre-Thorpe, M., & Bülthoff, H. H. (2001) Detection of  
984 animals in natural images using far peripheral vision. *European Journal of Neuroscience*,  
985 14, 869-876.

986 Williams, A. M., & Davids, K. (1998). Visual search strategy, selective attention, and exper-  
987 tise in soccer. *Research Quarterly for Exercise and Sport*, 69, 111-128.

988 Yantis, S. (1992). Multi element visual tracking: Attention and perceptual organization. *Co-*  
989 *gnitive Psychology*, 24, 295-340.

990 Yu, H. H., Chaplin, T. A., & Rosa, M. G. P. (2015). Representation of central and peripheral  
991 vision in the primate cerebral cortex: Insights from studies of the marmoset brain. *Neuro-*  
992 *science Research*, 93, 47-61.

993 Zelinsky, G. J., & Neider, M. (2008). An eye movement analysis of multiple object tracking  
994 in a realistic environment. *Visual Cognition*, 16, 553-566.

995 Zelinsky, G. J., & Todor, A. (2010). The role of “rescue saccades” in tracking objects through  
996 occlusions. *Journal of Vision*, 10, 1-13.

- 1  
2  
3 997 Zhong, S-H., Ma, Z., Wilson, C., Liu, Y., & Flombaum, J. I. (2014). Why do people appear  
4  
5 998 not to extrapolate trajectories during multiple object tracking? A computational investiga-  
6  
7 999 tion. *Journal of Vision*, 14, 1–30.  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

For Review Only